Chapter #9, FROM CONCEPTION TO DESIGN, A Practical Guide to Designing Ambient Displays

Jennifer Mankoff, and Anind Dey

IRB-TR-03-025

July, 2003

In Ohara, K. and Churchill, E. (eds.) Public and Situated Displays. Kluwer.

Intel Research Berkeley

Chapter #9

FROM CONCEPTION TO DESIGN

A practical guide to designing ambient displays

Jennifer Mankoff[1] & Anind K. Dey[1,2]
[1]EECS Department, UC Berkeley, Berkeley, CA 94720-1770

[2]Intel Research Berkeley, Berkeley, CA 94704-1307

Abstract: This chapter discusses displays that sit on the periphery of a user's attention. Many public displays of information that we encounter are in this category the majority of the time, including clocks, posters, and windows. Computationally enhanced variations on this theme are called *peripheral displays*. Our work focuses on *ambient displays*, a subset of peripheral displays that continuously display information to be monitored. Peripheral (and ambient) displays have the peculiar property that they are *not* meant to be the focus of the user's attention. Contrast this with desktop applications, which require a user's attention. In this chapter, we present a case study of two public ambient displays that we developed and evaluated. We present some lessons learned about the design of ambient displays, and conclude with a practical guide to using a modified version of heuristic evaluation that we developed as a result of designing these displays.

Key words: ambient displays, peripheral displays, heuristic evaluation, attention, iterative design

1. INTRODUCTION

Many of the public displays of information that we encounter sit on the periphery of a user's attention the majority of the time, including clocks, posters, and windows. Computationally enhanced variations on this theme are called *peripheral displays*. This chapter is concerned with *ambient displays*, a subset of peripheral displays that continuously display information that can be monitored by the user without requiring her focused attention. For example, one of the first such displays, created by an artist and technologist in collaboration, was a "dangling string" attached to a motor (Weiser and Brown, 1995). The string spun around at different speeds

depending on network load (See Figure #9-1). Contrast this with a discrete display of information such as an alarm, an *alerting display* that only rings when the network load reaches a certain threshold.

Figure #9-1. "The 'Dangling String' is an 8 foot piece of plastic spaghetti that hangs from a small electric motor hanging in the ceiling…. A very busy network causes a madly whirling string with a characteristic noise [left]; a quiet network causes only a small twitch every few seconds [right]. Placed in an unused corner of a hallway, the long string is visible and audible from many offices without being obtrusive" (Weiser and Brown, 1995). Courtesy of the Palo Alto Research Center (PARC). Photos by B. Tramontana.

Peripheral displays have the peculiar property that they are *not* meant to be the focus of a user's attention. They can be broken into two categories: *Ambient displays*, are a subclass of peripheral displays that present information in such a way so as to allow people to monitor a data source while not being distracted from their main task. *Alerting displays* alert a user through more direct means about salient information. This chapter will focus in particular on ambient displays. However, many alerting displays include an ambient component when they are not actively alerting the user about something, and thus alerting display designers may benefit from some of the material in this chapter. Similarly, an ambient display may at times alert a user about something.

One of the biggest challenges facing the developers of ambient displays is the lack of information about what constitutes a *good* design, and which ambient display designs will succeed or fail. Considering the literature regarding ambient displays, there have been few in-depth evaluations, and even fewer examples of design iteration (exceptions are the evaluation work of Mamyinka *et al.* (2001), and the design iterations of Mynatt *et al.* (2001)). Far more evaluations of alerting displays exist (including Cutrell *et al.*

(2001); Dumais and Czerwinski (2001); Ho-Ching *et al*. (2003); Horvitz *et al*. (1999); Maglio and Campbell (2000); Tiernan, *et al*. (2001) and van Dantzich *et al*. (2002)). In these cases, as in the ambient display case, discount evaluation techniques[1], useful in the early stages of design, do not exist.

The goal of this chapter is to provide a guide to ambient display design and evaluation. After reviewing relevant literature, we present our experiences with the design and evaluation of two ambient displays. We present metrics for successful ambient display design (changes in *awareness* or *behavior*), and describe lessons learned designing our displays. We then present an evaluation technique that we adapted explicitly for ambient display design, giving the reader a practical guide for how to apply it. We conclude with a discussion of future work in the design and evaluation of ambient displays.

2. BACKGROUND

2.1 Peripheral Displays

Peripheral displays are devices that usually reside in the *periphery*, but show information relevant to a user (Examples may be found in: Dumais and Czerwinski, 2001; Mamyinka *et al.*, 2001; McCrickard *et al.*, 2001; McCrickard *et al.*, 2002; Mynatt *et al.*, 1998; Mynatt *et al.*, 2001). By definition, in today's world of ubiquitous devices, the majority of applications must reside in the user's periphery—a person can only interact closely with one or two things at a time. As Weiser states in his seminal article on Ubiquitous Computing "Most of the computers… will be invisible in fact as well as metaphor (Weiser, 1991)" – not at the center of a user's task or attention. Treismann's theory of feature integration states that singular focused attention (or spotlight) is a necessity when interacting with complex information (Treismann and Gelade, 1980). In addition, other studies have shown that recognizing a change in a visual environment does not require this spotlight, although knowing what the change is does require it (Sagi and Julesz, 1985). Good design is crucial for applications, or displays of information, that reside in the periphery to succeed without representing a constant distraction, or further overloading users who are already overwhelmed by information in today's world.

[1] Discount evaluation techniques are techniques that have a reduced cost. They typically require less time and/or money, may involve fewer participants, and less extravagant equipment (Nielsen, 1994).

Peripheral displays may be public or personal displays of information, and may show ambient information, or may show alerts. An example of a personal, alerting display is Yahoo's Instant Messenger program. We classify it as an *alerting display* because it moves from the periphery to the focus of the user's attention to notify the user that someone wishes to communicate a message to her. Another example in this category is the mobile phone, which grabs the user's attention when it rings. In contrast Mamyinka *et al.* (2001) designed a personal, ambient display that indicates to a presenter using PowerPoint™ how much time is left in his presentation. We classify this as an *ambient display* because it presents information continuously to the presenter without interrupting the presentation in any way. An example of a public, ambient display is the Dangling String, described in the introduction and shown in Figure #9-1 (Weiser and Brown, 1995). Both of these ambient displays are abstract (one represents time as a color bar, the other represents network load as the rotation speed of a string. This is typical of computer-based ambient displays. A focus on aesthetics is also typical: The Dangling String is a piece of artwork, created to explore the idea of embodied information.

In summary, work in peripheral displays, whether personal or public, falls into two classes – ambient displays and alerting displays. Ambient displays are peripheral displays that do not attempt to move into the foreground of a user's attention. They are *continuous* displays of information, and are often abstract, aesthetic, and non-disruptive. In contrast to ambient displays, alerting displays show information at *discrete* intervals, and are defined by their attempts to move into the foreground of a user's attention. It should be noted that in practice, many peripheral displays include components of both ambient displays and alerting displays. The focus of this chapter is on continuous displays of information, ambient displays, although some of our displays may include alerting features.

2.2 Ambient Displays

Many ambient displays have been built over the years, and those presented here are only a sample. Ambient displays have become an increasingly important focus for artists, designers, and technologists since the introduction of the Dangling String. For example, the Oxygen Flute (Chafe and Niemeyer, 2002) was an art piece designed to reflect the amount of oxygen available in an enclosed space depending on the presence and breathing patterns of human participants. Traditional user interface design had little relevance to this provocative exhibit.

In contrast, Hallnäs and Redström’s (2001) work with slow technology was a deliberate exploration of the uses of slowly changing physical artifacts

such as moving curtains, wind-blown leaves, and drawers, to show (write) and sense (read) information. As well as being aesthetic and compelling, this work also begins to explore ideas that might guide future design.

Researchers approaching the problem from a human-computer interaction perspective have performed studies of detection, distraction, and other factors that can inform ambient display design. For example, Bartram *et al.* (2001) found that motion is an effective way to notify a user of information, while color change is less effective. McCrickard *et al.* (2003) studied a variety of different styles of animation and their impact in terms of interruption and comprehension. The result of their studies was a table recommending different types of notifications[2] depending on system design goals. Czerwinski *et al.* (2000) investigated the impact of interruptions caused by Instant Messaging on different types of primary tasks, showing generally harmful effects on overall task time. In related work, the same authors found that Instant Messaging interruptions can cause a user to forget what they were scanning for in a visual scan task (Cutrell *et al.,* 2001). Intille (2002) demonstrated techniques for supporting change blindness, changing visual information in ways that make the changes difficult to detect, and reduces interruption/distraction. Hudson *et al.* studied predictors for the interruptibility of administrative assistants, and developed sensing technology that can accurately predict appropriate times to display peripheral information, based on sensors as simple as a noise-detecting microphone. Other work, including our own, has led to an understanding of how different evaluation methods can be adapted to support the evaluation of peripheral and ambient displays (Mankoff *et al.*, 2003; Chewar and McCrickard, 2002).

Finally, technologists such as Ishii *et al.* (1998) and Pederson and Sokolor (1997) have developed tools and techniques for constructing tangible, sensing-based ambient displays. Both sets of authors used their tools to gain experience with ambient display design. Ishii *et al.* particularly mention the importance of choosing an appropriate modality to display data ("mapping"). They also discuss the importance of understanding when data will transition into a user's foreground, and the existence of learning effects. Pederson and Sokolor discuss the connection between learning effects and levels of abstraction. Long term use showed them the importance of aesthetics as they became "tired of the abstract displays" they had created. Finally, they suggest that because a user is not constantly attending to a display, information cannot change too quickly, and some sort of memory or history must be provided.

[2] Note that the peripheral display community has not entirely settled on a terminology yet. The term "notification systems" has been used in the past to refer to alerting displays, but McCrickard uses it more broadly and includes ambient displays in his definition.

One important question to ask about the evaluation of ambient displays is: "What are the characteristics of a successful ambient display?" In other words, what high-level outcomes should we expect from a good display? Our answer to this question is that a successful ambient display should modify someone's awareness of certain information and potentially change one's behavior with respect to that information. The next section presents our design process for two ambient displays, from conception to an evaluation intended to measure successful changes in awareness and behavior.

3. OUR DESIGN PROCESS

Our design process typically involves three main steps: (1) We conduct interviews and/or surveys of the people using the area where the display will be located to determine which information sources are of interest to them, and how they currently monitor that information. (2) We select an appropriate information source based on that data and design an ambient display that shows that information. (3) We conduct a summative evaluation of the display. Our general approach has been to test awareness and behavior of the information displayed twice – once before and once after the deployment. Next, we give an example of this design process for two recent displays built in our lab, the Bus Mobile and the Light Display.

Figure #9-2. Two displays that were placed at the front of an underground student computing laboratory. **Bus Mobile** (left): The three images show a token for bus 51 moving upwards, indicating that the bus is approaching the bus stop. **Light Display** (right): The light dims as sunset approaches, and brightens as dawn approaches.

3.1.1 Bus Mobile and Light Display

Both the Bus Mobile and the Light Display were built to address the needs of undergraduate computer science students working in underground laboratories in our building. We designed them to provide a service to the users of two computer labs that do not have windows.

To determine what kinds of information would be useful to users, we conducted a preliminary survey, asking participants to rank in order information they would like to see in an ambient display. Some of the choices included how dark it is outside (leaving the lab before dark is important to many students for safety reasons), the weather, the population of the computer labs, the network load on the login servers, when a bus was next scheduled to arrive, and sports statistics. The survey also queried the lab users with questions such as which bus lines they used frequently and which sports teams they followed. Some of the highest ranking information included how full the labs were and which servers had the most traffic. However, it was decided that these data sources would not make effective ambient displays, because the information only need be obtained once before entering the lab, and continual awareness of the information is not needed while one is working. Instead, we chose the highest ranked information sources that included dynamic change (and thus were amenable to an ambient display approach). The two most popular information sources that fit this requirement were bus schedules and the amount of daylight remaining. Students typically memorize bus schedules, or look them up online. Daylight is estimated by checking the time. As students are usually engrossed in their projects, it is very easy for them to lose track of time and more importantly, when their buses arrive or when it gets dark.

Once we completed our analysis of what data to show, we built the displays and deployed them in the aforementioned labs for a two week period. It was expected that having the displays in the two computer labs would allow the lab users to make an informed decision to leave the lab within a specific time frame based on their needs. For the Bus Mobile, we hypothesized that there would be an increase in the number of students who leave the lab within an optimal time frame before a given bus arrival, because students would like to catch a bus immediately instead of waiting at the stop for one to arrive. For the Daylight Display, we hypothesized that there would be an increase in the number of students leaving the lab in the time leading up to sunset.

Bus Mobile: The Bus Mobile, shown in Figure #9-2 (left), was designed to provide information about how soon local buses would arrive at nearby bus stops. It was constructed using six computer-controlled motors attached to aluminum rods. Around the base of the display hangs a curtain that is

approximately six inches long, behind which the icons, attached to each of the motors with a wire, can be hidden from view. Bus numbers painted on wooden plaques for the six most popular bus lines mentioned on the pre-survey were hung from the motors.

The display was designed so that the six bus numbers hang at depths corresponding to how many minutes are left before the relevant bus is scheduled to arrive at the stop closest to the building where the labs are located. For example, in Figure #9-2 (left), Bus 51N is further from the bus stop than Bus 51S. For each inch the number hangs below the bottom edge of the skirt, there is one minute remaining before the bus arrives, with a maximum range of twenty-five minutes. If no bus is scheduled to arrive within twenty-five minutes, the bus number moves to its minimum depth and is hidden from view behind the curtain. The display is updated every minute, and the icons move up one inch per minute. After the bus has passed, the bus number is lowered the appropriate number of inches based upon the next scheduled arrival of the same bus line. As mentioned above, we queried students to determine the most popular bus lines to display with the Bus Mobile.

Daylight Display: The second display, shown in Figure #9-2 (right), was designed to provide information about whether it was dusky, light, or dark outside. This was of interest to students for safety reasons (*i.e.* some students are uncomfortable walking home alone after dark) and also as an approximate indicator of time of day. The display consists of a regular floor lamp with an X10 controller, a common home automation device. The display controls the brightness of a lamp by adjusting how much power is supplied to the lamp.

The Daylight Display was designed to adjust its brightness according to the recorded sunrise and sunset times. A number of sunrises and sunsets were observed to determine when it actually begins to get light and dark, so that the display could more accurately show the light level outside. It was determined that during the summer months in Berkeley in 2002, the sun began to set about forty-five minutes before the listed sunset time (taken from a website with weather information), and it was completely dark about thirty minutes after the listed sunset time. For the sunrise, it began to get light thirty minutes before the listed sunrise time, and it was completely light about forty-five minutes after the listed sunrise time. The display was designed to alert a user of the lab that it is beginning to get dark or light outside by flickering the lamp a few times, then beginning the cycle of brightening or dimming the lamp, depending on the time of day. The X10 controller allowed for twenty-two brightness levels from off to on, which meant that the light adjusted its brightness approximately every 3.5 minutes.

3.1.2 Study Method

The study was conducted in three phases. The first phase occurred in the week before the displays were deployed in the labs, when a questionnaire was distributed to the users of three labs, which they could fill out at their own leisure. Two of the labs were ones where the displays would be located, and the third lab was one that would have no display and would be used as a control. The questions on the survey covered various topics that could give us an indication of whether or not information about bus arrivals and outdoor light levels affected a subject's behavior.

Phase 2 of the project started about a week and a half after the surveys were first made available. It began with the installation of the two displays in two different labs. The Bus Mobile was suspended from the ceiling on one side of the room in a location where most users in the lab could see it when they looked up from their desktop computers. The Daylight Display was set up in a corner of the other lab where it would be the most noticeable. We hung a sign by each of the displays, explaining how to use them and where a user could go for more information. As soon as the displays were deployed, we removed the Phase 1 surveys from the labs.

Phase 2 lasted for two and a half weeks. The first week of deployment allowed the users of the lab comfortable with using the ambient displays, so that when the next survey became available, the novelty of the displays would have worn off. Surveys with different questions were made available to the users of the labs, this time with questions pertaining to how they use the display and whether or not it was useful to them. Users had the opportunity to complete surveys posted on the door, using an online form with the web address advertised in the labs, or completing the survey after being recruited in the hallway after leaving.

Phase 3 started when the Bus Mobile and Daylight Display were removed, along with the surveys from the second phase. Surveys were again distributed to the labs in the same manner as the surveys from the second round. The objective of this questionnaire was to find data on the impact of the removal of the displays and whether or not the users of the labs missed having them.

In addition to the data collected from the surveys, we monitored the labs for ten minutes a day at random times for approximately two weeks, spread over phases 1 and 2. The observations made during the monitoring sessions consisted of how many people were in the lab initially, how many people entered or left the lab during the time period, how many users were there in the end, and then any observations that could be made about people using the displays, discussing them, or filling out the surveys. Our last source of

data was a system log that listed all of the students' account login and logout times. This data was retrieved in each of the phases.

3.1.3 Study Results

Qualitative feedback indicated that students preferred the Bus Mobile to the daylight display. For example, one student wrote "bus mobile -> ultra cool. makes life easier [sic]" about the Bus Mobile. Written comments criticized the Daylight Display for being too bright, or indicated that a respondent had not noticed it or thought it was broken.

We surveyed a total of 60 students during the study period. Of those, 6 respondents were interested in the bus schedules and 10 were interested in awareness of light (based on a Likert scale response of 4 or 5 out of 5). There was a strong correlation (Spearman's *rho*=.618, $p<.01$) between those who reported interest in the daylight information and those who reported finding the daylight display useful. There was a very strong correlation (*rho*=.808, $p<.01$) between interest in the bus schedules and the usefulness of the Bus Mobile, while there was a moderate correlation (*rho*=.595, $p<.01$) between interest in the bus schedules and respondents who missed it after it was removed. Although 83% of the respondents who were interested in the Bus Mobile found it useful, only 30% interested in the daylight found it useful.

As stated above, we observed the lab each day for ten minutes over the course of two weeks. During these observations, very few people left or entered the lab, and students did not attend to the displays in noticeable ways or discuss them with each other. Instead, they remained focused on their lab work during the entire observation period.

Lastly, we analyzed the login and logout data for the Bus Mobile. We removed certain data to make our results resemble the typical use of the lab. This included removing the entry of any user who was logged in for 5 minutes or less. This was to reduce the amount of data that might skew the results due to a user merely logging in to check mail or logging into a machine, then immediately deciding to use another one. Another set of confounding factors that could alter the results were the times that lab sessions were scheduled to meet in the rooms with the displays. All logout times just before, during, and just after all lab sessions were removed from the data set.

We then calculated a "logout delta" for each data point. The logout delta measured the time from a logout event to the most likely trigger for that event. To calculate the logout delta, we subtracted each logout time from the nearest bus arrival time.

To check for a statistical significance of the Bus Mobile on logout times, we ran a t-test across the three phases of the project. Our analysis was not able to show a statistically significant change in user logout behavior.

The above results, even without the additional login time data factored in, give some qualitative insight into whether or not the hypotheses were correct. For the prediction that the Bus Mobile and Daylight Display would impact people's decisions to leave the lab at a certain time, it can be said that the Bus Mobile was definitely effective for the people interested in the data, although not for as many people as originally expected. Some subjects had very positive comments about the Bus Mobile, saying they chose to work in the lab where it was displayed because they found it so convenient.

For the Daylight Display, the survey results show the display was only effective for about 30% of the subjects interested in the daylight schedule, which did not meet our expectations. There are several possible explanations for why this might be the case, using the comments from the surveys as an indication. For one, the design of the Daylight Display might not have been appealing. Many people saw it as just a lamp in the corner of the room and may not have noticed there was anything special about it unless they got closer to it and read the sign. Also, some subjects mentioned they never noticed the display working, which could mean a display that changes more frequently than two 75 minute long periods each day is more useful. One subject complained that the lamp was too bright and painful to the eyes, which would also make it less appealing. These are all things to take into consideration in our next iteration of the Daylight Display.

In summary, we used a user-centered design process to build two ambient displays and gain feedback on their usefulness and some of the problems with their design. Our analysis showed that these displays are useful to users interested in the information they display.

3.2 Lessons Learned

From a design perspective, we learned several things about ambient displays. First, we learned that it is crucial to pick the right sort of information source. Information that changes very rarely (such as whether it is light or dark out) is not appropriate for an ambient display, because it is difficult for the user to notice changes in the display. Users may get out of the habit of attending to it, since there is little change, and may receive little benefit from having the display. This issue interacts with the noticeability of the cue. For example, a window might succeed at the same goal because eventually a user will notice that he is having difficulty reading due to the fading of natural light.

Second, we learned that, since most of the people viewing a public ambient display may care nothing about the information it is displaying, a good design must be able to stand on its aesthetic merits alone. In other words, a display may be a purely decorative object to the majority of its "users." There may be some tension between this decorative role and the additional goal of displaying information, and this creates additional challenges for the display designer. It is also difficult to evaluate the aesthetic goals of a display, although examples of such evaluations exist (See Höök *et al.*, 2003).

Third, since those users who are interested in the data source still may not attend to a display much of the time, it must be able to smoothly transition from a background, peripheral object to the main focus of a user's attention. This transition may be difficult to manage, since the display cannot easily sense a user's interruptibility, or other aspects of her cognitive state. Although sensors may help with this issue (see, for example, the work of Hudson *et al.*, 2003), it also creates challenges for the designer. For example, a display may attempt to attract attention, demand attention, or simply allow a user to interact once it has a user's attention. Each of these possibilities bring about different sorts of design and interactivity challenges.

In retrospect, we felt that our designs would have benefited from more iterations before the actual deployment. As a result, we adapted a popular discount evaluation technique, *Heuristic Evaluation*, to the domain of ambient displays (Mankoff *et al.*, 2003). Heuristic evaluation does not require a working interface and thus can be used at the early stages of design. Additionally, it can be done by designers rather than end users, making it an inexpensive solution for supporting multiple design iterations. This is important since reports of development time for a single design iteration can be as high as a year (Heiner *et al.*, 1999). Below, we describe how our modified version of Heuristic Evaluation can be used to evaluate an ambient display.

4. USING HEURISTIC EVALUATION TO DEVELOP AMBIENT DISPLAYS

Heuristic evaluation involves recruiting evaluators, who may be novices, to critique an interface (usually represented with pictures and a textual description)[3]. Evaluators look for problems in an interface's compliance with heuristics that encode important usability guidelines. Nielsen found that 3-5

[3] Note that heuristic evaluation is most effective when performed by trained usability evaluators and does not help to identify which of the identified issues are real problems that will have a measurable impact on usability (Gray and Salzman, 1998).

novice evaluators find 40-60% of known issues when applying heuristic evaluation to desktop interfaces (Nielsen and Molich, 1990). The canonical list of heuristics as defined by Nielsen (2002) is: (1) Visibility of system status; (2) Match between system and the real world; (3) User control and freedom; (4) Consistency and standards; (5) Error prevention; (6) Recognition rather than recall; (7) Flexibility and efficiency of use; (8) Aesthetic and minimalist design; (9) Help users recognize, diagnose and recover from errors; (10) Help and documentation.

We believe heuristic evaluation is important to ambient display designers because of its potential to provide quick, inexpensive feedback about the possible issues with a specific display (Nielsen, 1990). Heuristic evaluation, because of its informal nature and low cost, was rated as one of the top techniques currently in use in a survey of usability practitioners (Vredenburg *et al*. 2002). Among discount evaluation techniques, only “informal expert review” ranked higher than heuristic evaluation. However, heuristic evaluation was considered to be more effective than informal expert reviews by the survey participants. Also, it was believed to be low cost, fast, and easy to apply. Cost and speed were both highly ranked factors in choosing an evaluation methodology.

The major difference between evaluating ambient displays and evaluating traditional displays comes from the way users interact with the interface. Ambient display users are passive in obtaining information from a display. Users do not *use* a display as they would use a computer; they *perceive* a display. Consequently, some of Nielsen's original heuristics (those associated with interactivity and productivity) do not apply to ambient displays. Furthermore, since ambient displays are passive, more responsibility lies with the designer to ensure that the displays are conveying a sufficient amount of information, and that the information is being conveyed unobtrusively. We were particularly concerned with how Nielsen's heuristics were defined, not only how they were named, as the definitions made assumptions about the goals of the interface being evaluated. For example, two heuristics explicitly refer to a “dialog” with which the user interacts, something that is central to conventional interfaces, but rare in ambient displays. Our conclusion was that the methodology of heuristic evaluation could be applied to ambient displays if the heuristics used were modified.

We adapted heuristic evaluation for use with ambient displays (Mankoff *et al*. 2003). With the help of ambient display designers, we defined and tested a modified set of heuristics. Next, we give a practical guide to conducting a heuristic evaluation of an ambient display using our modified heuristics.

4.1 How it Works

We begin by presenting the basic methodology, which we did not change from Nielsen and Molich's (1990) original description[4]. The participants in a heuristic evaluation are generally designers. They may or may not be domain experts. They should not be the same people who developed the original display. Three to five evaluators generally conduct a heuristic evaluation. Five evaluators are typically cited as sufficient to find 80% of issues when conducting a heuristic evaluation (Nielsen and Landauer 1993). However, Woolrych and Cockton (2002) criticize the claim that five users is sufficient on the basis that problems may not be evenly distributed by severity, and users may not consistently find 30% of problems (an assumption of Nielsen and Landauer's (1993) original claim). These criticisms are directed not at the technique itself, simply at the number of users necessary to conduct an effective evaluation.

4.1.1 Method

Each evaluator is given a description of the user interface to be evaluated, a list of the heuristics to be used, and a form in which she or he can enter problems as they are found. For each problem found, a description of the problem is written down and a severity is assigned to the problem. Evaluators work independently to generate problems.

Once each evaluator has generated a list of problems, the designer of the ambient display must organize the results. One or more designers work together to group problems that are similar. Cox and Greenberg (2000) have developed a method for doing this systematically using *results synthesis*. Designers may also reassign severity levels at this point.

Heuristic evaluation is a qualitative technique, since it cannot be used to measure how well an interface meets numerical metrics. Once the problems found by the participants have been coded, no further analysis need be done. At this point, most designers work through the interface addressing each problem in turn.

The heuristics given below are based on our previous work (Mankoff *et al.*, 2003), but have been modified slightly as follows: Originally we determined that complete problem coverage would require all but one of the heuristics we had derived, as well as five of Nielsen's original heuristics (Match between system and real world, Visibility of system status, User control and freedom, Error Prevention, and Flexibility and efficiency of use). However, Nielsen's heuristics, although effective, are not entirely

[4] We refine some details of this methodology that are underspecified, particularly in how to analyze the results produced by evaluators.

appropriate for ambient displays. Thus, we combined “User control and freedom” and “Error prevention”, because our evaluators used them both to capture problems with identifying the occurrence of an error. Additionally, we deleted the least relevant heuristic, “Flexibility and efficiency of use”, because it would be difficult or impossible to tailor a public display to the needs of specific, individual users. Table #9-1 has the full set of heuristics.

Table #9-1. Heuristics for use when conducting an Heuristic Evaluation of Ambient Displays. Heuristic (7) is a direct quote from Nielsen (2002)

#	**Heuristic and definition**
1.	**Sufficient information design:** The display should be designed to convey “just enough” information. Too much information cramps the display, and too little makes the display less useful.
2.	**Consistent and intuitive mapping:** Ambient displays should add minimal cognitive load. Cognitive load may be higher when users must remember what states or changes in the display mean. The display should be intuitive.
3.	**Match between system and real world:** The system should speak the user’s language, with words, phrases, and concepts familiar to the user. Follow real-world conventions, making information appear in a natural and logical order.
4.	**Visibility of state:** An ambient display should make the states of the system noticeable. The transition from one state to another should be easily perceptible
5.	**Aesthetic and pleasing design:** The display should be pleasing when it is placed in the intended setting.
6.	**Useful and relevant information:** The information should be useful and relevant to the users in the intended setting.
7.	**Visibility of system status:** The system should always keep users informed about what is going on, through appropriate feedback, within reasonable time.
8.	**Easy transition to more in-depth information:** If the display offers multi-leveled information, the display should make it easy and quick for users to find out more detailed information.
9.	**“Peripherality” of display:** The display should be unobtrusive and remain so unless it requires the user’s attention. Users should be able to easily monitor the display.
10.	**Error prevention and user control:** Users should be able to distinguish between a inactive display and a broken display. When a display enters an error state, or a state that a user does not want, users must have a way to reset it.

4.1.2 Special considerations for ambient displays

As stated above, our major modification to the heuristic evaluation technique was to generate and validate a new list of heuristics appropriate for ambient displays. We validated our heuristics by comparing their performance to Nielsen's heuristics on the two ambient displays described in Section 3. Evaluators using our heuristics found more, severe problems than evaluators using Nielsen's heuristics. Additionally, when using our heuristics, 3-5 evaluators were able to identify 40-60% of known usability issues. This implies that heuristic evaluation is an effective technique for

identifying usability issues with ambient displays, given our modified set of heuristics.

Our modified heuristics are given in Table #9-1, with definitions. Our work showed that even people with little experience designing ambient displays could be useful evaluators given our list of heuristics. We recommend selecting people with general design experience, however. One other detail is necessary for successful evaluation of an ambient display using heuristic evaluation: it is important to set the scene in which the ambient display is operating. A description of its location in a public space, information about what tasks might be co-occurring with the display, and so on, are all helpful to the evaluator. We are still working to determine what the necessary elements of such a description might be, but the high-level important message is that an ambient display cannot be understood in isolation, but must be contextualized for the evaluator. As an example, here is a description of the Bus Mobile, used in our validation of the new heuristics.

> "The Bus Mobile is a mobile that hangs from the ceiling and has six bus numbers hanging from it, each representing a distinct route, which adjust their height according to the closeness of a bus. It is designed to show when the next bus for various bus lines is scheduled to stop at the location closest to Soda Hall. The Bus Mobile is located in Soda Hall Instructional Lab 273. It is pretty much the typical computer lab setting with desks, chairs and computers. The Bus Mobile is hung from the ceiling in the front of the lab, near the left side of the whiteboard. All the desks and chairs are lined up facing the whiteboard. It is important to note that there are no windows to the outside in the lab. When a bus for a particular bus line is 25 minutes from the bus stop, its icon drops to its maximum depth. For each minute that passes, moving closer to the bus arrival time, the icon moves up towards the white skirt [which is visible at the top of Figure #9-2]. When the bus is not scheduled to arrive, the icon moves completely under the white skirt, disappearing from view and stays out of sight until 25 minutes before it is scheduled to arrive again. If buses for a particular bus line are scheduled to arrive less than 25 minutes apart, the icon's height represents the bus that will arrive next."

We also provided the evaluators with four time-delayed pictures of the 51N bus token in motion upwards towards the white skirt of the display.

4.2 Summary of Heuristic Evaluation

As stated above, we believe that heuristic evaluation is a viable design tool for constructing ambient displays. Our own studies proved that our

modified heuristics were effective for evaluating ambient displays (Mankoff *et al.*, 2003). The technique is simple and low-cost, and can help ambient display designers to improve their designs before running a more costly, longer-term evaluation.

5. CONCLUSIONS & FUTURE WORK

In conclusion, this chapter focuses on the design and early-stage evaluation of ambient displays, a subset of peripheral displays that help a user to maintain awareness of information by continuously displaying information in a way that can easily be monitored without requiring focused attention. In this chapter, we contrast ambient displays with alerting displays, which typically display information at discrete intervals, when it is important to gain a user's attention. In practice, peripheral displays may include properties of both.

As discussed in Section 2, a mixture of artists and technologists has built a variety of interesting ambient displays over the years since Jerimijenko, Weiser, and Brown first presented the Dangling String (Figure #9-1). However, much less is known about what makes a particular display succeed or fail, or how, exactly, to design for specific attentional issues.

We described a study of two such displays that we built, the Bus Mobile and Light Display, in Section 3. Although time consuming, the study lead to a set of lessons about ambient display design. In particular, we learned that the rate of information change, and the noticeability of change, are crucial to display success. We also learned that public displays create a tension between aesthetics and information content. This is because they will have many groups of users with varying interests. Third, we learned that a good design must carefully mediate between the importance of the information being displayed (which may change over time), and the amount of attention required. This is a design challenge. Last, we learned how important iterative design and early feedback are to display success.

In Section 4, we showed how a particular technique, Heuristic Evaluation, can be used to study ambient displays. This technique is ideal for early-stage iteration because it does not require a working display to succeed. Using the heuristics presented in this paper, around 5 evaluators can find many potential problems without the overhead of an extensive study such as that described in Section 3.

Evaluation of ambient displays is clearly difficult. However, we believe this is due in part to a lack of evaluation techniques that are appropriate to use during the early stages of design of ambient and peripheral displays. Our

modifications to heuristic evaluation have resulted in a practical, effective technique to use during these early stages.

Ideally, a designer should have a choice of more than one early stage technique, and our current and future efforts are focused on expanding the repertoire of techniques available. In particular, we are looking at the role of paper-based prototyping in ambient and peripheral display design.

We are also interested in exploring the usefulness of our heuristics in helping to develop design guidelines. However, we believe that peripheral display design requires more than qualitative information such as that encoded in our heuristics. We are currently investigating the literature on attention in cognitive science in the hopes of establishing a scientific basis for deciding between different manifestations of information.

Lastly, we are working to create a tool that can support rapid prototyping of peripheral displays. Such a tool should allow designers to easily experiment with different levels of noticeability and different types of transitions indicating changes in information. This would allow a display designer to focus on aesthetics while handling the underlying tasks necessary to a display changing computational information over time.

REFERENCES

Bartram, L., Ware, C., and Calvert, T. (2001). Moving icons: Detection and distraction. Proceedings of Human-Computer Interaction—Interact; 2001, July; Tokyo, Japan. IOS Press, IFIP.

Chafe, C., and Niemeyer, G. (2002). Oxygen flute. Installation by Artists, San Jose Museum of Art, October 13, 2001–June, 2002.

Chewar, C., and McCrickard, D. S. (2002). Adapting UEMs for notification systems. Design and evaluation of notification interfaces for ubiquitous computing, Ubicomp 2002: workshop 9; 2002, September 29-October 1; Göteborg, Sweden.

Cox, D. and Greenberg, S. (2000). Supporting collaborative interpretation in distributed groupwear. Proceedings of the CSCW 2000 Conference Computer Supported Collaborative Work; 2000, December 2-6; Philadelphia, PA. ACM Press. pp. 289–298.

Cutrell, E., Czerwinski, M. & Horvitz, E. (2001). Notification, disruption and memory: Effects of messaging interruptions on memory and performance. Proceedings of Human-Computer Interaction—Interact; 2001, July; Tokyo, Japan. IOS Press, IFIP. pp.263-269.

Czerwinski, M., Cutrell, E. & Horvitz, E. (2000). Instant Messaging and Interruption: Influence of Task Type on Performance. OZCHI 2000 Conference Proceedings; 2000, December 4-8; Sydney, Australia. pp. 356-361.

Dumais, S. and Czerwinski, M. (2001). Building bridges from theory to practice. Proceedings of HCI International 2001, 9th Conference on Human-Computer Interaction. 2001, August 5-10; New Orleans, LA, USA. Lawrence Erlbaum Associates. pp. 1358-1362.

Gray, W. D., and Salzman, M. C. (1998). Damaged merchandise? A review of experiments that compare usability evaluation methods. Human-Computer Interaction 1998; **13**(3): 203–261.

Hallnas, L., and Redstrom, J. (2001). Slow technology; designing for reflection. Personal and Ubiquitous Computing 2001; **5**(3):201–212.

Heiner, J. M., Hudson, S. E., and Tanaka, K. (1999). The information percolator: Ambient information display in a decorative object. Proceedings of the UIST 1999 Symposium on User Interface Software and Technology. 1999, November 7-10; Asheville, NC, USA. ACM Press. pp. 141–148.

Ho-Ching, W. L., Mankoff, J., and Landay, J. (2003). From data to display: the design and evaluation of a peripheral sound display for the deaf. Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. 2003, April 05-10; Ft. Lauderdale, CA, USA. ACM Press. pp. 161-168.

Höök, K., Sengers, P. and Andersson, G. (2003). Sense and sensibility: Evaluation and interactive art. Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. 2003, April 05-10; Ft. Lauderdale, CA, USA. ACM Press. pp. 241-248.

Horvitz, E., Jacobs, A., and Hovel, D. (1999). Attention-sensitive alerting. Proceedings of the UAI 1999 Conference on Uncertainty and Artificial Intelligence (UAI 99). 1999, July 30-Aug 1; San Francisco, CA, USA. Morgan Kaufmann. pp. 305-313.

Hudson, S.E., Fogarty, J., Atkeson, C.G., Forlizzi, J., Kielser, S., Lee, J.C., and Yang, J (2003). Predicting human interruptibility with sensors: A Wizard of Oz feasibility study. Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. 2003, April 05-10; Ft. Lauderdale, CA, USA. ACM Press. pp. 257-264.

Intille, S. S. (2002). Change blind information display for ubiquitous computing environments. Proceedings of Ubicomp 2002, LCNS 2498. 2002, September 29-October 1; Göteborg, Sweden. Springer-Verlag. pp. 91-106.

Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., and Yarin, P. (1998) Ambient displays: Turning architectural space into an interface between people and digital information. Proceedings of the CoBuild 1998 First International Workshop on Cooperative Buildings, LCNS 1370. 1998, February; Darmstadt, Germany. Springer-Verlag. pp. 22-32.

Maglio, P.P. & Campbell, C.S. (2000). Tradeoffs in displaying peripheral information. Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems. 2000, April 1-6; The Hague, The Netherlands. ACM Press. pp. 241-248.

Mamykina, O., Mynatt, E.D., and Terry, M. (2001). Time aura: Interfaces for pacing. Proceedings of the CHI 2001 Conference on Human Factors in Computing Systems. 2001, March 31-April 5; Seattle, WA, USA. ACM Press. pp. 144-151.

Mankoff, J., Dey, A., Hsieh, G., Kientz, J., Lederer, S., and Ames, M. (2003). Heuristic evaluation of ambient displays. Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. 2003, April 05-10; Ft. Lauderdale, CA, USA. ACM Press. pp. 169-176.

McCrickard, D. S., Catrambone, R. and Stasko, J. T. (2001). Evaluating animation in the periphery as a mechanism for maintaining awareness. Proceedings of Human-Computer Interaction—Interact; 2001, July; Tokyo, Japan. IOS Press, IFIP. pp. 148-156.

McCrickard, D. S., Catrambone, R., Chewar, C. M. and Stasko, J. T. (2003). Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. International Journal of Human-Computer Studies 2003; **8**(5):547-582.

McCrickard, D.S., Wrighton, D., and Bussert, D. (2002). Supporting the construction of real world interfaces. Proceedings of the 2000 IEEE Symposia on Human Centric Computing Languages and Environments. 2002, September 03-06; Arlington, VA, USA. IEEE Press. pp. 54-56.

Mynatt, E., Back, M., Want, R., Baer, M., and Ellis, J. B. (1998). Designing Audio Aura. Proceedings of the CHI 1998 Conference on Human Factors in Computing Systems. 1998, April 18-23; Los Angeles, CA. ACM Press. pp. 566-573.

Mynatt, E., Rowan, J., Craighill, S., and Jacobs, A. (2001). Digital family portraits: Providing peace of mind for extended family members. Proceedings of the CHI 2001 Conference on Human Factors in Computing Systems. 2001, March 31-April 5; Seattle, WA, USA. ACM Press. pp. 333–340.

Nielsen, J. (1994). "Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier." In *Cost-Justifying Usability,* R. G. Bias and D. J. Mayhew, ed., Academic Press, 1994. pp. 245-272.
(Online Version: http://www.useit.com/papers/guerrilla_hci.html)

Nielsen, J. (2002) Ten usability heuristics. Retrieved August, 2002 from: http://www.useit.com/papers/heuristic/heuristic_list.html

Nielsen, J. and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. Proceedings of the INTERCHI 1993 Conference on Human Factors in Computing Systems. 1993, April 24-29; Amsterdam, The Netherlands. ACM Press. pp. 206-213.

Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces. Proceedings of the CHI 1990 Conference on Human Factors in Computing Systems. 1990, April 1-5; Seattle, WA, USA. ACM Press. pp. 249–256.

Pederson, E.R. and Sokoler, T. (1997). Aroma: abstract representation of presence supporting mutual awareness. Proceedings of the CHI 1997 Conference on Human Factors in Computing Systems. 1997, March 22-27; Atlanta, GA. ACM Press. pp. 51-58.

Sagi, D., and Julesz, B. "Where" and "what" in vision. Science 1985; **228**:1217-1219.

Tiernan, S. L., Cutrell, E., and Czerwinski, M. (2001). Effective notification systems depend on user trust. Proceedings of Human-Computer Interaction—Interact; 2001, July; Tokyo, Japan. IOS Press, IFIP. pp. 684-685.

Treismann, A., and Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology 1980; **12**:97-136.

van Dantzich, M., Robbins, D., Horvitz, E., and Czerwinski, M. (2002). Scope: Providing awareness of multiple notifications at a glance. Proceedings of the AVI 2002 6th International Working Conference on Advanced Visual Interfaces. 2002, May 22-24; Trento, Italy. ACM Press.

Vredenburg, K., Mao, J.-Y., Smith, P. W., and Carey, T. (2002). A survey of user-centered design practice. Proceedings of the CHI 2002 Conference on Human Factors in Computing Systems. 2002, April 20-25; Minneapolis, MN, USA. ACM Press. pp. 471–478.

Weiser, M. (1991). The computer for the 21st century. Scientific American 1991; **265**(3):94–104.

Weiser, M., and Brown, J. S. (1995). Designing calm technology. PowerGrid Journal. 1996; **1**(1). (Online Version: http://www.ubiq.com/weiser/calmtech/calmtech.htm).